Journal of Surgical Research 171, e81–e95 (2011)
doi:10.1016/j.jss.2011.06.034

RESEARCH REVIEW

Methods and Tools for Objective Assessment of Psychomotor Skills in Laparoscopic Surgery

Ignacio Oropesa, B.Sc.,*,†,1 Patricia Sánchez-González, M.Sc.,*,† Pablo Lamata, Ph.D.,*
Magdalena K. Chmarra, Ph.D.,‡ José B. Pagador, B.Sc.,§ Juan A. Sánchez-Margallo, B.Sc.,§
Francisco M. Sánchez-Margallo, Ph.D.,§ and Enrique J. Gómez, Ph.D.*,†

*Bioengineering and Telemedicine Centre (GBT), ETSI Telecomunicación, Universidad Politécnica de Madrid (UPM), Madrid, Spain; †Networking Research Center on Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Zaragoza, Spain; ‡Department of BioMechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering (3mE), Delft University of Technology, Delft, The Netherlands; and §Minimally Invasive Surgery Centre Jesús Usón, Cáceres, Spain

Originally submitted March 4, 2011; accepted for publication June 15, 2011

Training and assessment paradigms for laparoscopic surgical skills are evolving from traditional mentor–trainee tutorship towards structured, more objective and safer programs. Accreditation of surgeons requires reaching a consensus on metrics and tasks used to assess surgeons' psychomotor skills. Ongoing development of tracking systems and software solutions has allowed for the expansion of novel training and assessment means in laparoscopy. The current challenge is to adapt and include these systems within training programs, and to exploit their possibilities for evaluation purposes. This paper describes the state of the art in research on measuring and assessing psychomotor laparoscopic skills. It gives an overview on tracking systems as well as on metrics and advanced statistical and machine learning techniques employed for evaluation purposes. The later ones have a potential to be used as an aid in deciding on the surgical competence level, which is an important aspect when accreditation of the surgeons in particular, and patient safety in general, are considered. The prospective of these methods and tools make them complementary means for surgical assessment of motor skills, especially in the early stages of training. Successful examples such as the Fundamentals of Laparoscopic Surgery should help drive a paradigm change to structured curricula based on objective parameters. These may improve the accreditation of new surgeons, as well as optimize their already overloaded training schedules.

Key Words: **laparoscopic surgery; objective evaluation; accreditation; metrics; basic skills; virtual reality; human motion tracking; classification.**

INTRODUCTION

Minimally invasive surgery (MIS) has changed the way surgery is performed in operating rooms (OR). In many procedures, it has become the recommended standard, displacing open surgery [1]. Laparoscopy, one of the most common MIS approaches, has been adopted by several surgical sub-specialties, including gastrointestinal, gynecologic, and urologic surgeries [2]. Effective training and assessment of surgeons in these new techniques have become one of the major concerns of hospitals and clinics in recent years, fuelled mostly by patients' and society's demands for safer surgeries and well prepared physicians [3–6].

Traditional Halsted-based training [7] is potentially unsafe for the patient and, as a consequence, no longer ethically sustainable [4]. There is a tendency to move the early phases of training, concerned with the acquisition of motor skills, outside the OR. For this reason, laboratory settings including, for example, box trainers and/or virtual reality (VR) simulators have been developed [4].

Another issue that has become evident is the need for structured formation programs inside the OR [8]. Assessment based on In-Training Evaluation Reports (ITERs) [9] is subjective, expensive, and prone to

[1] To whom correspondence and reprint requests should be addressed at Bioengineering and Telemedicine Centre (GBT), ETSI Telecomunicación, Universidad Politécnica de Madrid (UPM). Avda Complutense, 30, 28040, Madrid, Spain. E-mail: ioropesa@gbt.tfo.upm.es.

0022-4804/$36.00

undesirable side effects, such as a halo effect [10]. The halo effect is a result of influenced perception of the performance in one area (e.g., laparoscopic task) by the performance in another area (e.g., relationship between trainee and mentor) [11]. Additionally the ITERs are periodically written, depriving the trainee of immediate feedback, and are subject to the evaluators' long-term memory [12].

Structured reports, based on checklists and immediate end-product analysis have been proposed and validated [11,13–16]. The Objective Structured Clinical Examination (OSCE) [13] has been used to assess clinical performance of trainees on various clinical stations (comprising different tasks). Since OSCE mainly focuses on the assessment of procedural knowledge and attitude of the trainee towards the patient, technical evaluation of psychomotor skills is not given leading relevance [9].

The Objective Structured Assessment of Technical Skills (OSATS) pays more attention to motor skills assessment [14]. Validity has been fully established for skills assessment ranging from simple tasks to advanced, complex procedures [15]. Implementation of the OSATS in the OR, however, may present ambiguities in the scoring systems [16]. Another drawback is the high amount of resources required, from the number of experts deployed at each station to evaluate the trainees, to the marginal costs of each exam per candidate [9]. Laparoscopic video offline-evaluation has been proposed to reduce some of these costs with good reliability results [17]. However, the presence of a reviewer is still required, and trainees do not obtain immediate feedback about their performance. A counterpart of OSATS for MIS—the Global Operative Assessment of Laparoscopic Skills (GOALS)—has been developed by Vassiliou *et al.* [11]. GOALS are not procedure-specific reports, and as such, they can be used for any MIS procedure.

As the need for objective and structured assessment of technical performance grows, new tools and methods have emerged over the last few years [9, 15, 18]. Training methods are being gradually changed, leaving the traditional ways behind on behalf of criterion-based curricula [19]. This tendency has been favored by the development and advances on tracking systems and computing technologies, which have led to the appearance of human motion tracking devices and virtual simulation for surgical training [20]. With new ways to measure surgical performance, the remaining question is: "What does it mean to be a competent surgeon?" Thus, simultaneous to the technological advances, much research has been devoted to the development of metrics for skills assessment. These metrics determine to a great extent a measuring instrument's proficiency, and are necessary to provide evidence of its reliability and validity as an assessment tool [8] (Table 1).

The purpose of this review is to present a state-of-the-art on the new tools available for acquisition and analysis of information concerning surgical performance, and their influence in the development of new accreditation programs. For this end, we have modeled the process of surgical assessment as a three-sided problem: (1) the clinical side; (2) the technological side; and (3) the analytical side [21]. The clinical side deals with the definition of the optimal tasks, metrics and conditions to consider for the assessment of the different psychomotor skills required. The technological side is related to the use of tracking technologies and/or computer assisted systems for the creation of surgical training and assessment environments that allow capturing objective data concerning a surgeon's skill. Finally, the analytical side studies the use of statistical analysis and machine learning algorithms for data analysis, in order to ascertain whether automatic classification systems to aide surgical assessment are viable or not.

METHODOLOGY

Search of the literature was performed using PubMed and Google Scholar public databases. Key words employed were: "laparoscopy", "minimally invasive surgery," "surgical assessment", "psychomotor skills", "objective evaluation", "validation", "virtual reality, " and "motion analysis." Obtained article' bibliographies were also checked for new references. Additionally, validation brochures for commercial virtual reality simulators were scanned. For all considered sections, articles not related with laparoscopic surgery were filtered. No *a priori* restrictions regarding publishing date or language were applied.

For the Clinical Definition of Metrics section, recovered results were scrutinized and filtered by construct validation studies employing objective data. Our purpose was to cross-reference the most recurrent metrics and surgical tasks for box trainers or VR simulators, identifying which parameters influence on the different psychomotor skills. In this way, it was possible to discern general patterns where a given metric/set of metrics yields significant differences between training groups, with respect to different tasks and abilities. Thus, only reports showing positive results are included in this review. A complete up-to-date detailed revision on valid/nonvalid systems is presented in [18].

In the Tracking Technologies for Skills Assessment section, a technological overview of tracking tools for skill assessment is given, from an application point of view. Depending on the setting on which the tracking systems are used, we propose to categorize them as those used in (1) VR simulated environments, and (2) real settings. The first one refers to tracking